MTJ-Based Nonvolatile 9T SRAM Cell

S. Kushwaha, D. Kumar, M. Saw, A. Islam
Department of Electronics and Communication Engineering
Birla Institute of Technology, Mesra, Ranchi, Jharkhand, India
dhrhuva1157.09@bitmesra.ac.in, monisha1216.09@bitmesra.ac.in

Abstract—This paper presents a spin-transfer torque- magnetic tunnel junction (STT-MTJ) based non-volatile 9-transistor (9T) SRAM cell. The cell achieves low power dissipation due to its series connected MTJ elements and read buffer which offer stacking effect. The paper studies the impact of PVT (process, voltage, and temperature) variations on the design metric of the SRAM cell such as write delay and compares the results with non-volatile 8T SRAM cell (NV8T). The proposed design consumes lower leakage power and exhibits narrower spread in write delay compared with NV8T.

Index Terms—variability, SRAM, RSNM, WSNM, MTJ, hold power

I. INTRODUCTION

Spintronics is the field of electronics that exploit the spin property of electrons giving rise to spin transistors; contrary to charge of electrons in traditional electronic devices [1-7]. It gives a new direction for research where these spin transistors can be combined with ordinary transistors to extract the maximum potential of both. The promising features of spin transistors are non-volatility and re-configurability. These features are very useful and suitable for future on-chip, high-speed, high-density, low-power and nonvolatile memory circuits. The nonvolatile property of spin transistor is best suited for power-gating systems and field programmable gate arrays [8].

There are two kinds of magneto-resistive devices – spin valve (SV) and magnetic tunnel junction (MTJ). Spin valve (SV) devices consist of at least two ferromagnetic metal (FM) layers separated by an ultra-thin nonmagnetic metal (NM) layer. In an MTJ, two ferromagnetic layers – one free layer and the other fixed layer are separated by an oxide layer. The insulating oxide layer is of either AlOx or MgO. The MgO is now preferred to AlOx due to its higher tunnel magneto-resistance (TMR) ratio defined as (R_{AP} - R_{AP})/R_{AP}, where R_{AP} and R_{AP} are the resistances of MTJ in anti-parallel and parallel magnetic orientation [9]. Spin-transfer torque (STT) phenomenon is used to change the magnetic orientation of the free layer relative to the fixed layer to give a parallel and an antiparallel direction. The parallel orientation of MTJ offers low resistance and conversely, an antiparallel orientation offers high resistance. These two resistance states can be used to denote a logic state of “0” and “1” respectively. These resistance states persist even after the removal of switching current passing through the MTJs. This property is utilized in nonvolatile memory design. The MTJ is also CMOS-compatible with high stability and reliability [10]. Moreover, the STT-MTJ can be fabricated on top of CMOS devices to reduce the area overhead [11]. All these features raise hope for building a nonvolatile memory and logic circuit that do not consume OFF-state leakage current and support ultra-low power operation. This work utilizes MTJ in conjunction with conventional MOSFET to mitigate variation in design metrics of memory circuit such as 9-transistor SRAM cell [12].

The memory has evolved into a multi-leveled structure with SRAM making up the fastest access memory. In modern microprocessors, this trend has continued, such that SRAM cache itself has multiple levels. For instance, in recent Intel and AMD microprocessors, L1 (level one) cache contains a small amount of the fastest memory cells. At the next level, L2 cache contains a large amount of cells that are slightly slower, and so on. 6T- SRAMs are widely used in L1 and L2 cache arrays while DRAMs are used as large off-chip memory arrays or as embedded DRAMs (eDRAM) as higher level, high-density caches. Conversely, the design of traditional charge-based SRAMs, DRAMs and eDRAMs pose challenges such as higher standby power and lower noise margin in sub-45 nanometer nodes and beyond. Alternative to 6T-SRAM bit-cells, subthreshold 10T [13], low-power 10T [14], and low-leakage 11T [15] SRAM cells have been proposed for low-voltage and low-power operations. However, the problem of standby power in semiconductor memory remains to be solved.

CMOS technology scaling driven by the benefit of integration density, higher speed of operation and lower power dissipation, has overcome many barriers over the past few decades. Presently, it is facing even more complications, which are more severe than earlier. One of them is variability. Variability is becoming a metric of equal importance as power, delay, and area. TG-based fully differential 8T SRAM bitcell proposed in [16] could mitigate variability in design metrics of SRAM cell, but failed to improve read static noise margin (RSNM). This work attempts to mitigate variability in design metrics of MTJ-STT-based nonvolatile memory cell with enhanced read static noise margin (RSNM) in highly scaled technology node such as 22-nm. The non-volatility of the proposed design solves the problem of leakage power in memory. Extensive simulations were performed in HSPICE using 22-nm Predictive Technology Model (PTM) [17] to support the concluding remarks.

The rest of the paper is organized as follows. Section II presents the proposed design. Simulation results are discussed and compared in Section III. Finally, the concluding remarks are provided in Section IV.
II. PROPOSED DESIGN AND DEVICE SIZING

This paper proposes a STT-MTJ-based nonvolatile differential memory cell (see Fig. 1) (hereafter called NV9T). Its design metrics are assessed and compared with its differential counterpart namely STT-MTJ-based nonvolatile 8T SRAM cell (see Fig. 2) (hereafter called NV8T). The transistors in NV9T cell was sized to maintain pull-up (PU) transistors < pass-gate (PG) transistors < pull-down (PD) transistors. Transistor widths for NV9T cell $W_W/4W_P/2W_D$ were chosen as $22n/33n/44n$ respectively. For a fair comparison, transistors of basic 6T portion of both the cells had the same dimensions. Widths of the extra transistors in NV9T (MN5/6/7), were equal to $33n/33n/44n$ respectively. For extra transistors MN5/6 in NV8T a width of 44 nm was selected. Channel length of all transistors of both the cells was taken to be 22 nm. $\alpha_{\text{ratio}}$ and $\beta_{\text{ratio}}$ of 6T portion of both the cells were taken to be 1.33 and 0.67 respectively, where $\alpha_{\text{ratio}} = \beta_{\text{drive}}/\beta_{\text{access}}$ and $\beta_{\text{ratio}} = \beta_{\text{pull-up}}/\beta_{\text{access}}$. This is because typical values of $\alpha_{\text{ratio}}$ ranges from 1.2 to 3 to avoid read upset in conventional 6T SRAM cell [18]. As write-ability of the SRAM cell is determined by the $\alpha_{\text{ratio}}$ or pull-up ratio, the $\beta_{\text{ratio}}$ was decided to be $\beta_{\text{ratio}} = 1.8$ to maintain good write-ability [19]. Therefore, to maintain an appreciable read stability, write-ability and feedback pull-up strength $\beta_{\text{ratio}} = 1.33$ and $\beta_{\text{ratio}} = 0.67$ were chosen for the proposed design [13]–[16].

The critical design strategy of the proposed cell was the series connection of an MTJ with each driver. These MTJs help in storing and restoring the cell content. A read buffer (MN5/6/7) was used for the cell to isolate storage nodes while reading. This helped in achieving higher read stability. The proposed cell would exhibit very small RSNM without this read buffer since the drivers (MN1/2) were very small. Channel length of all transistors of both the cells had the same dimensions. Widths of the extra transistors in NV9T (MN5/6/7), were equal to $33n/33n/44n$ respectively. For extra transistors MN5/6 in NV8T a width of 44 nm was selected. Channel length of all transistors of both the cells was taken to be 22 nm. $\alpha_{\text{ratio}}$ and $\beta_{\text{ratio}}$ of 6T portion of both the cells were taken to be 1.33 and 0.67 respectively, where $\alpha_{\text{ratio}} = \beta_{\text{drive}}/\beta_{\text{access}}$ and $\beta_{\text{ratio}} = \beta_{\text{pull-up}}/\beta_{\text{access}}$. This is because typical values of $\alpha_{\text{ratio}}$ ranges from 1.2 to 3 to avoid read upset in conventional 6T SRAM cell [18]. As write-ability of the SRAM cell is determined by the $\alpha_{\text{ratio}}$ or pull-up ratio, the $\beta_{\text{ratio}}$ was decided to be $\beta_{\text{ratio}} = 1.8$ to maintain good write-ability [19]. Therefore, to maintain an appreciable read stability, write-ability and feedback pull-up strength $\beta_{\text{ratio}} = 1.33$ and $\beta_{\text{ratio}} = 0.67$ were chosen for the proposed design [13]–[16].

The critical design strategy of the proposed cell was the series connection of an MTJ with each driver. These MTJs help in storing and restoring the cell content. A read buffer (MN5/6/7) was used for the cell to isolate storage nodes while reading. This helped in achieving higher read stability. The proposed cell would exhibit very small RSNM without this read buffer since the drivers (MN1/2) were very small and MTJs were in series with them. Thus, this read buffer improved the read stability of the proposed cell.

A. Simulation Setup

This section presents comparison of various design metrics, which were estimated during MC (Monte Carlo) simulation using 22-nm PTM. In order to analyze variability with MC simulations, channel length (L), channel-doping concentration (NDEP), oxide thickness ($t_{\text{ox}}$) and threshold voltage ($V_T$) of all the devices in the cells were considered as independent random variables and assumed to have Gaussian distribution with $3\sigma$ variation of 10% [21].

Anticipated fluctuation in $V_{DD}$ is 10% in future technology generations such as 22-nm [22]. Therefore, most of the design metrics were estimated by varying the supply voltage by ±10% around the nominal $V_{DD}$ of 0.8 V. Authors in [23] showed that a sample size of 2000 ensures a lower than 4% inaccuracy in the estimation of standard deviation. Hence, a sample size of 5000 [24] was used in this work to achieve even higher accuracy. All parameters were estimated using the above simulation setup.

B. Store and Restore Operation

In normal SRAM operation of NV8T SR and CTRL signals remain grounded. Storage operation in NV8T is initiated by raising SR and CTRL lines to $V_{DD}$ (virtual $V_{DD}$). At the end of storage operation, SR and CTRL pulses are turned low. Then global $V_{DD}$ and GND are disconnected from the $V_{DD}$ and $V_{GND}$ (virtual GND) by sleep transistors (see Fig. 3) and MRAM array enters into inactive mode in which no read/write operations are performed.

Values of the storage nodes are stored in MTJ in the form of $R_{AP}/R_{P}$ (anti-parallel resistance/parallel resistance) in the MTJ. During restore operation, SR and CTRL lines remain grounded and $V_{DD}$ is raised high by switching sleep transistors ON. Store and restore operation of proposed NV9T SRAM cell is illustrated in Fig. 4. During the store operation,
$V_H$ drops due to discharging of storage node “H” through the leaky devices. When the restore operation is initiated by switching the sleep transistor ON the storage node “H” is restored to its previous high level.

C. Read Access Time or Read Delay

Read delay is estimated as the time required by BL (bit line)/BLB (bit line bar) to discharge by 50 mV [24]. The 50 mV differential between BL and BLB is good enough to be detected by sense amplifier, thereby avoiding misread [24], [25]. The proposed cell exhibits 38% penalty in read delay at nominal $V_{DD}$ compared with NV8T (see Fig. 5). This penalty occurs due to larger body effect offered by the read buffer and larger bit line capacitance.

D. Write Access Time and Its Variability

Write delay is estimated as the time required by the node storing “0” to rise up to 90% of $V_{DD}$ or the node storing “1” to fall by 10% of $V_{DD}$. As pull-up and access devices are mainly responsible for the write delay and they are of same sizes in both the cells, therefore, only marginal differences in the write delay is observed due to tail MTJ elements in NV9T compared with NV8T (see Table I and Fig. 6). However, the proposed design proves its robustness by exhibiting narrower spread (6% @ nominal $V_{DD}$ of 0.8) in write delay variability.

E. Write-ability

The write-ability of a cell is measured by write static noise margin (WSNM), which is estimated graphically using read and write VTCs (voltage transfer curves). The write VTC, while writing “1” to “L”, is gauged by sweeping $V_H$ (Fig. 7, y-axis) with BLB & WWL high and BL low and monitoring $V_L$ (Fig. 7, x-axis). This write VTC is used in combination with the read VTC, which is measured by sweeping $V_H$ (Fig. 7, x-axis) and monitoring $V_L$ (Fig. 7, y-axis). The side length of the smallest square, which can be embedded between the read and write VTCs of the same SRAM cell at the lower half of the curves, represents WSNM [26].

When WSNM falls below zero, write VTC intersects read VTC, thereby indicating positive write margin and signifying write failure. As can be observed from Fig. 7, the proposed design NV9T exhibits equal write-ability to NV8T. The proposed bitcell exhibits negative write margin, indicating successful write operation. As the proposed cell exhibits higher WSNM (=340 mV) than RSNM (= 330 mV), the cell is RSNM limited (see Fig. 8).

F. Read Stability

The standard 6T SRAM cell is found to be unstable at...
nanoscale technology generation. It and its variants fail to meet operational requirements due to a low RSNM. The RSNM is estimated graphically as the length of a side of the largest possible square that can be inscribed inside the smaller wing of the butterfly curve. Fig. 8 plots the “butterfly curves” for RSNM of NV9T and NV8T. As can be observed from Fig. 8, NV9T offers higher RSNM compared with NV8T (1.94×) due to isolation of storage nodes from bit lines. Cells with RSNM of at least 25% of $V_{DD}$ are generally considered to have excellent read stability [27]. The proposed cell meets this requirement.

G. Performance Analysis

For the estimation of the operating frequency, the precharge periods was not taken into account because it can be completely overlapped with the decoder operation. The estimated results are presented in Table II. As observed from the table, the proposed design exhibits lower operating frequency compared with NV8T because of larger body effect and bit line capacitance.

**H. Leakage Power Dissipation**

In nanoscale technology, thinner oxide and lower $V_t$ result in significant rise in gate leakage and sub-threshold leakage currents. Therefore, leakage power is now a significant contributor to the total chip power. It is imperative to reduce leakage power to prolong battery life. The comparison of hold power ($P_{HH}$) of NV9T and NV8T is shown in Fig. 9 and 10. As can be observed, the proposed design consumes lower leakage power at all temperatures and supply voltages. This is due to the larger body effect offered by read buffer and the use of MTJ elements at the source end of the pull-down drivers.
CONCLUSIONS

This work proposes a low-power and variability-aware nonvolatile 9T SRAM cell. It analyzes the impact of PVT variations on write delay. It consumes lower standby power at all considered temperatures and supply voltages. It also achieves higher read stability (RSNM) without reduction of its write-ability. The proposed design is, therefore, an attractive alternative for low power applications in scaled technology node such as 22-nm and beyond.

REFERENCES